

基于多尺度层级 LSTM 网络的时间序列预测分析

张旭东¹ 杜家浩¹ 黄宇方¹ 石东贤² 缪永伟³

(浙江工业大学计算机科学与技术学院 杭州 310023)¹ (浙江经贸职业技术学院 杭州 310018)²

(浙江理工大学信息工程学院 杭州 310018)³

摘要 现有的深度学习研究都依赖于网络的自发学习能力,在训练过程中力求避免或尽量减少人为先验知识的设定,导致网络训练过程完全“黑盒”,研究人员很难从语义上进行阐述。针对这种情况,文中提出了一种基于原始 LSTM 网络的改进——多尺度层级 LSTM(Multi-Scale Hierarchical Long Short-Term Memory, MSH-LSTM)网络。该网络保留了神经网络的常规实现流程,在网络学习过程中将层级网络结构与人的经验知识有机结合,使网络在人为指引下有针对性地训练,不再是完全的“黑盒”,同时对时间序列更好地进行分析预测。为说明 MSH-LSTM 网络结构的有效性,实验选取了两种时间序列数据(气温、股票),结果表明,相较于 ANN 网络、LSTM 网络及 GRU 网络,MSH-LSTM 网络在保证网络适用性的同时更具分析预测优势。在气温实验中,由于 MSH-LSTM 与常规 LSTM,GRU 网络都利用了序列数据的时间因素,因此,它们的效果明显优于 ANN;在股票实验中,MSH-LSTM 的 MAPE 误差相对于常规 LSTM,GRU,ANN 网络分别平均提升了约 19.65%,24.35%,46.30%。

关键词 LSTM,时间序列,短期预测,循环神经网络,层级网络

中图法分类号 TP391 **文献标识码** A

Time Series Analysis Based on MSH-LSTM

ZHANG Xu-dong¹ DU Jia-hao¹ HUANG Yu-fang¹ SHI Dong-xian² MIAO Yong-wei³

(College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China)¹

(ZheJiang Institute of Economics and Trade, Hangzhou 310018, China)²

(College of Information Engineering, Zhejiang Sci-Tech University, Hangzhou 310018, China)³

Abstract Nowadays, most researches in deep learning depend on the self-learning capacity of used neural network. Specifically, they focus on using as less human-knowledge priors as possible during the training step, which leads to totally “black-box” and is hard to clarify the training process semantically for researchers. In light of this situation, this paper proposes an improved structure of primitive LSTM (Multi-Scale Hierarchical Long Short-Term Memory, MSH-LSTM). It retains the common procedure widely used in neural network, combines the structure of neural network and human’s prior knowledge, enables the network to train purposefully under the guidance and solving the problem of “black-box” in a way, ultimately resulting in much better analytic results on time series data. To illustrate the effectiveness of MSH-LSTM, two groups of experiments (temperature and stock-price respectively) were carried out. Experimental results demonstrate that the proposed MSH-LSTM outperforms primitive ANN, LSTM and GRU without loss of network’s applicability. In temperature experiment, MSH-LSTM, primitive LSTM and primitive GRU use temporal information to get approximately results, which are better than primitive ANN. In stock price experiment, MSH-LSTM’s superiority is more obvious. The error of MAPE of MSH-LSTM is increased by an average of 19.65%, 24.35%, 46.3% compared with that of primitive LSTM, GRU and ANN, respectively.

Keywords LSTM, Time series, Short-term forecast, Recursive neural network, Hierarchical network

1 引言

时间序列是一组按照时间先后顺序排列且内部关联的数据。通过对历史时间序列的分析,能在一定程度上达到对未来长期或短期的精确预测,进而为生产实践创造更大的收益。截至目前,人们对时间序列的分析研究工作已经扩展到了气

象、经济、医学、电力、交通等诸多领域,针对时间序列进行分析,具有广阔的应用前景和巨大的现实意义。

早期用于时间序列分析的模型主要集中于传统线性模型,该分析模型简单、计算速度快,但假定了未来值与过去值、随机误差之间存在一定的线性函数关系,因此常局限于平稳时间序列的研究^[1]。随着研究逐步深入,线性模型不再满足

人们的准确度要求,且无法适用于实际复杂的应用场景,于是出现了非线性模型分析方法,如神经网络、支持向量回归、遗传规划算法。非线性模型具有较强的学习和数据处理能力,不需要事先假设数据之间具有何种函数关系,通过对大量样本数据的训练,模型能够自发地逼近复杂甚至很难用数学式描述的非线性特征。

对于常规的人工神经网络,样本输入节点之间相互离散,即网络本身并不关注输入节点的时序关系,但这种关系在某些特定问题中往往不可或缺,比如文本情感分析^[2]、污染物监测^[3-4]、医学跟踪^[5]、价格预测^[6]等问题的分析与前面若干时刻的输入密切相关,同样的数据在不同输入时序下会得到完全不同的结果。为了解决以上问题,人们提出了循环神经网络(RNN),该网络结构引入了输入数据的时间先后顺序,并使得分析结果得到显著提升。Saad^[6]对 RNN、时延网络与概率网络进行了实验比较;Lin 等^[7]首先利用卷积网络来提取数据的高维特征,然后借助 RNN 网络挖掘数据内部的时序特性,该组合网络改进了分类效果,同时提供了更好的模型解释性。

不过,传统 RNN 网络普遍存在着梯度消失、梯度爆炸的问题,很大程度上限制了长时序预测效果,于是 Hochreiter 等^[8]提出了长短时记忆网络(LSTM,即带 LSTM 记忆单元的 RNN 神经网络)来针对性解决该问题,使分析结果有了较大提升,并吸引了大量研究人员研究相关模型的发展及优化。Zhang 等^[9]和 Yang 等^[10]均用 LSTM 网络来预测海平面温度(SST)。其中,Zhang 等^[9]用 LSTM 层挖掘数据内部的时序联系,然后通过全连接层将 LSTM 层的输出映射到最终预测结果;Yang 等^[10]提出的网络结构包含 LSTM 层和卷积层,在利用数据时间信息的同时利用了空间信息,从而提高了预测精度。Karim 等^[11]提出了 ALSTM-FCN 网络结构,在简化数据集预处理步骤的同时,提升了分类效果。Zhang 等^[4]结合了小波变换和 LSTM 网络,首先针对变化剧烈的原始时间序列数据采取小波变换操作,得到多个变化相对较缓的子序列,然后分别对子序列建立 LSTM 网络,最终的预测结果是各 LSTM 网络预测结果的和。Liu 等^[12]的 ANR-SAE-VALSTM 模型先对原始时间序列数据做去噪声处理,再从去噪数据中提取特征并作为 VALSTM 网络输入,生成预测结果,在北京 PM2.5 及 GEFCom2014 上的实验表明了该网络结构的有效性。Zhang 等^[13]提出 LSTM 网络与 SVM 组合结构来预测电路故障,LSTM 从大规模的电力传输数据中提取时间特征,SVM 因其出色的鲁棒性负责生成最终的分类结果。还有很多研究者致力于对 LSTM 进行改进,如 Cho 等^[14]提出了 GRU,减少了记忆单元中门的数量,在保证原有效果的同时减少了参数量,随后 Chen 等^[15]提出了一种用 GRU 实现的基于社交平台评论来预测未来股市波动的方法;Sun 等^[16]从 LSTM 与 GRU 中获得启发,提出用 GMU 记忆单元网络完成手写体识别任务。

本文采用的多尺度层级 LSTM 是一种对传统 LSTM 的优化方案,该想法源自文献^[2,17],但与其应用场景不同,它

们研究的问题局限在分类问题;同时 Shi 等^[2]的多层次结构通过横向增加输入信息特征来提高分类准确率,没有考虑从低维数据到高维表示过程中纵向的信息丢失。本文将层级网络思想应用到了回归问题,并提出了 MSH-LSTM 网络结构,实验结果表明,该结构在时间序列预测分析问题中具有有效性。

2 LSTM 网络

LSTM(Long Short-Term Memory)神经网络是一种带有长短时记忆单元的 RNN 改进网络,能够很好地解决常规 RNN 中存在的梯度消失或梯度爆炸问题^[10]。每个记忆单元引入了 3 个门,分别为输入门、遗忘门、输出门,从而在网络训练中能够实现根据数据内部的联系自动控制各门的开闭和记忆单元值的更新,并赋予网络可变的“记忆”长度^[8]。输入门用来更新单元值 c ;遗忘门用来决定上一时刻的单元值 c 被保留的比例;输出门生成隐层状态值 h ,用作下一时刻的额外输入^[9]。LSTM 单元结构如图 1 所示

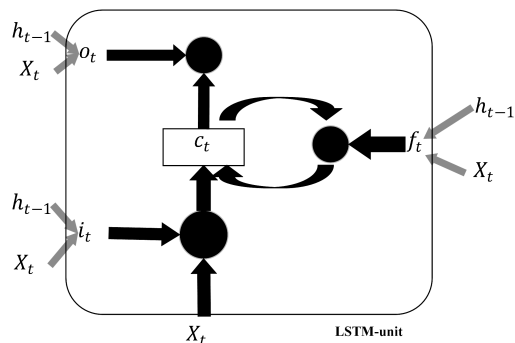


图 1 LSTM 单元结构^[8]

各门的控制信号和记忆单元值的更新公式^[18]如下:

$$\begin{cases} i_t = \sigma(W_i X_t + V_i h_{t-1} + b_i) \\ f_t = \sigma(W_f X_t + V_f h_{t-1} + b_f) \\ o_t = \sigma(W_o X_t + V_o h_{t-1} + b_o) \\ c_t = f_t \odot c_{t-1} + i_t + \tanh \odot (W_c X_{t-1} + V_c h_{t-1} + b_c) \\ h_t = o_t \odot \tanh(c_t) \end{cases}$$

在 t 时刻,LSTM 记忆单元接收 X_t ,同时与上一时刻($t-1$ 时刻)的状态输出 h_{t-1} 作为 LSTM 单元的输入;随后,输入与各门的 X 项权值 W 、 h 项权值 V 、偏置项 b 结合,并在 $\sigma(\cdot)$ (门激活函数)作用下,产生分别控制输入门(i_t)、遗忘门(f_t)、输出门(o_t)的信号;根据当前时刻各门的信号,更新记忆单元 c_t 以及生成当前时刻的状态输出 h_t ,并将其用作后续 $t+1$ 时刻的额外输入^[19],如此反复。训练过程中,LSTM 网络能够自发地学习到各门最佳的权值 W 、 V 以及偏置值 b ,使得单元值 c 在时间维度上能自发调节“记忆”时长。相较于常规的 RNN 神经网络,LSTM 网络能够很好地处理梯度消失、梯度爆炸问题^[8],提高模型精度,在近几年中已被广泛地应用于自然语言处理、文本情感分析^[2]、语义分析等问题。

3 MSH-LSTM 网络

目前针对神经网络的研究工作都力图最大程度地发挥网

络自发学习的能力,训练过程力求避免或尽量减少人为预先给定先验知识,导致整个网络的训练过程完全是一个黑盒。虽然这种做法极大减少了人工消耗,并在很多实验中取得了优于传统算法的时间序列分析效果,但在某种程度上,也导致分析效果的好坏完全取决于网络自身,同时由于训练过程完全“黑盒”,研究人员难以在语义上对训练过程做出合理的阐述,进而加大了后续网络结构调整、参数优化的难度。

鉴于以上问题,本文提出了一种多尺度层级 LSTM 网络。网络结构示意图如图 2 所示。

本文主要贡献如下:

(1)网络学习能力与人为给定的先验有机结合,使网络训练过程在人为引导下进行。

(2)研究发现,很多时间序列数据,如历史气温、地震波形、股市波动、文本等,都呈现出一定的层级规律。譬如,气温数据逐年反复、逐季度呈现爬升及下降趋势、短时趋于稳定;一篇文章可大致分为段落、句子、字词的多级结构,而文章的情感分析离不开上下文联系,无法从单个层级精确分析文章情感。如果能够利用数据中蕴含的此类内部层级规律,势必会改善分析效果,因此本文引入了层级网络结构,支持网络从多尺度对历史时间序列数据进行学习,以提高分析预测效果。

(3)层级网络学习过程中,下层到上层的正向传播过程存在着信息的丢失。MSH-LSTM 网络结构通过补偿丢失信息来保证时间序列分析取得更好的效果。

下面主要介绍 MSH-LSTM 网络实验处理流程,其与大多数网络操作流程一致:预处理、数据分片、网络模型训练、结果分析^[1]。

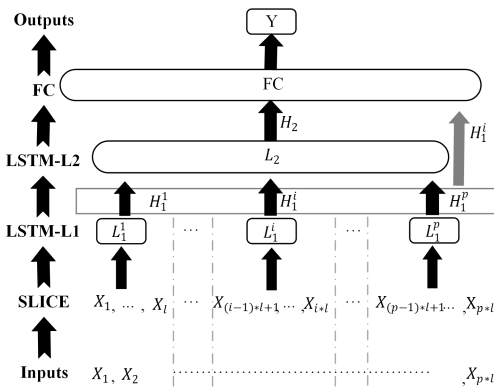


图 2 MSH-LSTM 网络结构示意图

3.1 数据预处理

数据的预处理主要包括完整性和归一化步骤。时间序列数据的完整性,用于确保每一时刻监测值的有效性、真实性,是开展后续研究工作的前提;数据的归一化处理,用于消除数据不同量纲之间的影响,处理后将数据限制在一定范围内,以便加速网络训练收敛和提高分析预测的精度。

对于短时平稳的原始时间序列数据,完整性处理主要采用均值填充方法来填补缺省值,具体公式为:

$$x_t = \frac{x_{t-1} + x_{t+1}}{2} \quad (1)$$

其中, x_t 表示待填补的 t 时刻数据值; x_{t-1} 和 x_{t+1} 分别表示前一时刻和后一时刻的数据值。接着,需要对时间序列数据进行归一化处理,将值标准化到统一的数值区间。使用的归一化公式如下:

$$X_m = \frac{X_{\text{origin}} - \min}{\max - \min} \quad (2)$$

其中, X_{origin} , \min , \max 分别为原始数值、训练数据中最小值、最大值, X_m 为归一化后的值,最后得到的 $X_i (i \in \{1, 2, 3, \dots\})$ 都被归一化到 $[0, 1]$ 区间,保障网络训练过程能够较快收敛。

3.2 数据分片

经过数据归一化处理后,得到标准的数据集合 X_m , 其中 $m \in R^+$ 。对于数据集中的训练样本,本文采用结合人为经验知识的方式,按指定的步长(用 l 表示)将样本输入数据分割成多个片段,记作 S_n , $S_n = \{X_j | (n-1) * l + 1 \leq j \leq n * l\}$, $n \in \{1, 2, 3, \dots, p\}$, 其中 p 为经验给出的分片数,分片得到的 p 个片段作为 MSH-LSTM 网络的输入。

3.3 MSH-LSTM 模型训练

区别于常规网络,MSH-LSTM 网络模型引入了层级网络思想,实现了从多个尺度对原始序列数据进行分析。该网络主要由两个 LSTM 网络层(L1, L2)以及一个全连接(FC)层组成, L1 层从低维时间层面探讨数据的时序关系; L2 层基于 L1 层的状态输出学习内含的高维时间关系;最后通过一个 FC 层给出整个 MSH-LSTM 网络输出,即网络的预测结果。

LSTM-L1 网络层以分片后各个片段作为输入,每个片段分别对应一个 LSTM 网络(记作 L_i , $i \in \{1, 2, \dots, p\}$), 其中下标 1 表示当前为 LSTM-L1 层,上标 i 表示第 i 个分片对应的 LSTM 网络。同时,该层各网络之间互不干扰,输出为网络的隐层状态 H_1^i , 其表示与 L_i^i 类似。

LSTM-L2 网络层以 L1 层各子网络隐层状态 H_1^i 作为输入,即将 L_1^1 子网络的输出 H_1^1 作为该网络的第一时刻的输入,将 L_2^2 子网络的输出 H_1^2 作为第二时刻的输入,如此反复; L2 层的输出结果表示为 H_2 。

FC 层以 H_2 作为输入,同时为了克服 L2 网络正向传播过程中出现的数据信息丢失问题,本文再将 L1 的输出结果 H_1^i 作为额外的输入项,从而能够提高 MSH-LSTM 网络在多尺度分析过程中数据的可靠性,保障学习效果。网络最终的输出结果记作 Y 。

3.4 过拟合

损失函数为网络训练效果的参考项。网络训练过程,可被看作是不断调整参数以使损失函数最小化的过程。假设样本量为 N , 通过网络正向传播得到输出 y^* , 而期望输出为 y , 则可以给出如下的损失函数^[3]:

$$\text{loss} = \frac{1}{N} \sum_{i=1}^N (y_i^* - y_i)^2 \quad (3)$$

然而,仅仅使用这样的损失函数最终很可能导致过拟合问题,即对训练集数据效果很好,而对实际的测试集数据效果很差。因此,本文采用了带正则化项的损失函数^[20], 即

$$loss = \frac{1}{N} \sum_{i=1}^N (y_i^* - y_i)^2 + \frac{\lambda}{2} \sum_w \omega^2 \quad (4)$$

其中, ω 为网络中各项参数, λ 为正则化项的权重。

3.5 评价指标

为量化实验结果便于比较, 本文选取了常用的平均绝对百分比误差^[9] (MAPE), 其公式如下:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left(\frac{|y_i^* - y_i|}{|y_i|} \times 10^{-6} \right) \quad (5)$$

其中, n 表示预测样本量, y^* 为预测值, y 为真实值。式中减去 1×10^{-6} 是为了防止出现真实值为 0 的情况。

4 实验分析

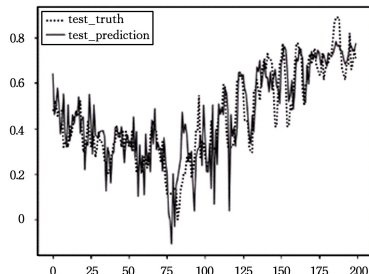
4.1 网络参数设置

为综合分析 MSH-LSTM 神经网络的分析效果, 将其与常规 ANN 网络、LSTM 网络、GRU 网络进行了比较, 每个网络的参数详见表 1。

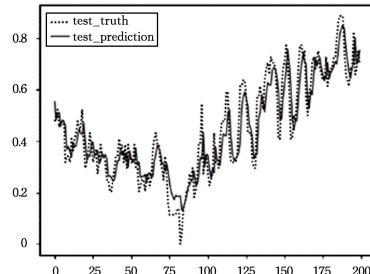
表 1 网络参数设置

网络结构	L1 节点数	L2 节点数	FC 层节点数	学习率	优化器	激活函数	迭代
ANN	10	10	—	0.01	Adagrad	tanh	200
LSTM	10	10	10	0.01	Adagrad	tanh	200
GRU	10	10	10	0.01	Adagrad	tanh	200
MSH-LSTM	10	20	10	0.01	Adagrad	tanh	200

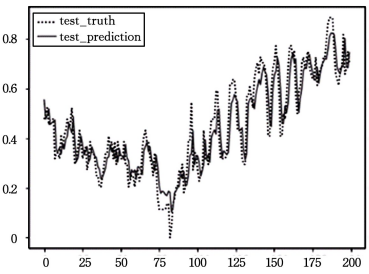
同时为说明 MSH-LSTM 网络的适用性, 本文选取了两组数据集(杭州历史日最高气温数据、股票数据)进行实验。



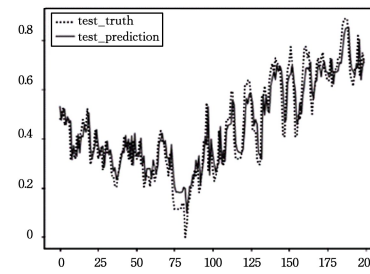
(a) 原始 ANN



(b) 原始 LSTM



(c) 原始 GRU



(d) MSH-LSTM

图 3 杭州气温分析预测部分截图

表 2 MAPE 统计

	原始 ANN	原始 LSTM	原始 GRU	MSH-LSTM
气温实验	0.148946736	0.107879142	0.106586844	0.105022057

4.3 股票实验

4.3.1 数据集说明及分片

实验随机抽取了两只股票的历史日收盘数据, 分别为 SH600030(中信证券)以及 SH601628(中国人寿), 时间分别

4.2 气温实验

4.2.1 数据集说明及分片

实验选取 2011-01-01—2018-05-29 七年余的杭州历史日最高气温数据。其中, 以 2011-01-01 到 2016-12-31 这六年数据(约占 80%)作为训练集, 以 2017-01-01 至 2018-05-29 的气温数据(约占 20%)作为测试集。样本以类似滑动窗口方式产生, 样本输入为前 21 天日最高气温, 输出为第 22 天日最高气温。MSH-LSTM 网络中分片设定为 3 片(人为指定的经验值, 对于不同研究问题可能有所不同), 因此, LSTM-L1 层可被看作研究一周内逐日气温的变化关系, 而 LSTM-L2 层探究三周中, 逐周的温度变化关系, 从而从两个尺度(周、日)分析历史气温变化, 实现对未来一天的最高气温做出更精确的预测。

4.2.2 实验结果

气温实验部分结果如图 3 所示, 图中点线表示真实值, 灰色实线表示网络预测值(后续实验结果图表示相同)。由于利用了序列顺序关系, MSH-LSTM 网络与原始 LSTM, GRU 网络取得了近似的效果, 对剧烈变动值有更好的把握, 而原始 ANN 的分析结果显得不可信, 无法对未来值进行可靠地预测。同样, MAPE 误差统计(见表 2)表明, 受益于层级网络多尺度(周、日)分析结构, 网络更好地把握了数据之间的时序变化关系, 从而提高网络的学习效果, 得到更精确的预测结果。由于气温数据集相对比较平稳, 因此 MSH-LSTM 网络结构较原始 LSTM 及 GRU 网络的优势并不明显, 在波动更为剧烈的股票数据中, 其优势更明显。

从 2003-01-06 和 2007-01-09 开始到 2018-03-19。为方便处理, 本实验训练集的选取都从起始日期至 2016-12-31, 并以 2017-01-01 到 2018-03-19 的数据作为测试集。样本的组织形式与气温实验类似, 选取的样本输入为历史 15 天的日收盘价, MSH-LSTM 网络分片仍为 3 片。

4.3.2 实验结果

股票实验部分结果如图 4 所示。

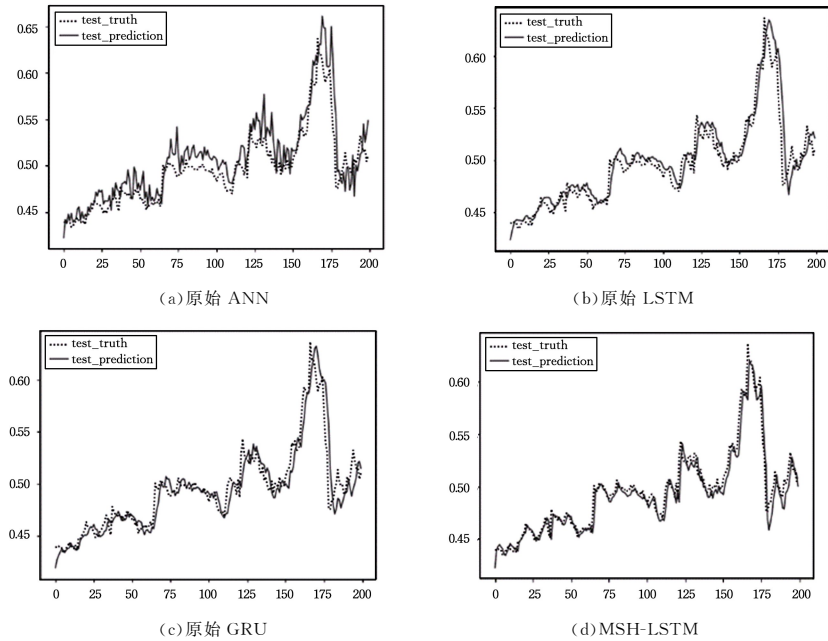


图4 SH600030[中信证券]分析预测部分截图

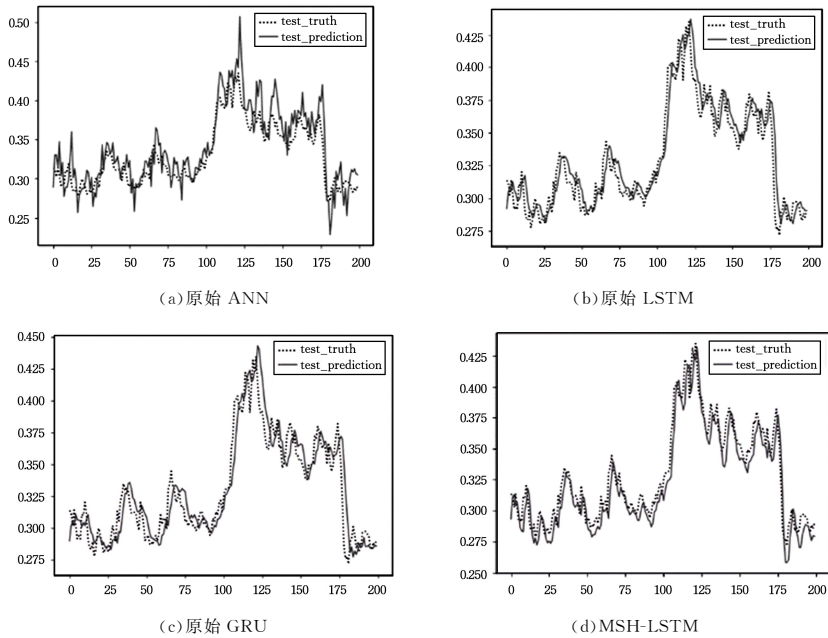


图5 SH601628[中国人寿]分析预测部分截图

比较图4、图5可以发现,MSH-LSTM预测结果明显优于其他3种原始网络。针对股票收盘价的预测,分析发现通过三周(交易日)的历史收盘波动预测未来一天收盘价要更为可靠,因此本实验同样取分片数目为3。综合比较4种网络实验结果,在随机选取的两支股票中MSH-LSTM都表现出了更高的精度和更强的鲁棒性,能够很好地预测未来可能出现的突变值;而原始LSTM及原始GRU虽能大体上把握股票收盘的变动趋势,但在剧烈波动的值域范围中效果欠佳;4种网络结构中,原始ANN效果最不理想,其最主要的一个原因是,ANN网络的样本输入之间是离散的值,不考虑输入之间的时序变动关系,因此对于股票价格这种复杂、影响因素众多、时序关系明显的场景,往往分析预测效果较差。4种网络在股票分析实验中的优劣从表3的误差统计中也能体现出来。

表3 MAPE统计

股票	原始 ANN	原始 LSTM	原始 GRU	MSH-LSTM
SH600030	0.029606607	0.016092418	0.017981726	0.012933817
SH601628	0.477055630	0.378437054	0.382682177	0.303848216

4.4 实验分析

比较误差统计(见表2、表3)可以清晰地看到,MSH-LSTM的分析预测效果要优于其余3种算法。在场景相对简单的气温实验中,MSH-LSTM网络结构相较于原始LSTM、原始GRU网络优势较小,主要表现在对小范围剧烈变动气温的预测,不过与原始ANN网络相比还是有明显的优势;在变动剧烈的股票实验中,MSH-LSTM预测的优势更为突出。以股票SH600030实验为例,其MAPE相比原始LSTM提升了近19.6%,比原始GRU提升了近28.1%,比原始ANN提升了近56.3%;同时,股票SH601628实验中,本文算法误差

分别降低了 7.5% (较 LSTM)、7.9% (较 GRU)、17.3% (较 ANN), 相对原始 LSTM、原始 GRU、原始 ANN 分别提升了 19.7%、20.6%、36.3%。

结束语 文中提出了一种 MSH-LSTM 网络结构, 有机地结合自身学习优势以及人的经验知识, 从而使得网络能够在人为指引下多尺度地从时间序列中发掘数据之间的内部关联。实验结果表明, 本文所提出的网络结构有效, 切实提高了分析预测效果。但 MSH-LSTM 网络也存在一些局限性:

1) 网络引入了人的经验知识, 势必会降低网络的适用性, 为了更好地分析预测效果, 往往需要针对问题人为地初步分析时间序列数据, 若不考虑对原始输入做分片处理 (即分片数 $p=1$), MSH-LSTM 网络可被看作普通的两层原始 LSTM 网络结构。

2) 网络增加了参数量, 因此需要更多的样本集进行训练, 计算开销加大, 同时较原始 LSTM、GRU 网络更难收敛, 整个网络的收敛效果受制于最差的子 LSTM 网络, 这点在更多层的网络结构中会更为明显。

参 考 文 献

- [1] SELVIN S, VINAYAKUMAR R, GOPALAKRISHNAN E A, et al. Stock price prediction using LSTM, RNN and CNN-sliding window model[C]//2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI). IEEE, 2017.
- [2] SHI S, ZHAO M, GUAN J, et al. A hierarchical lstm model with multiple features for sentiment analysis of sina weibo texts[C]//International Conference on Asian Language Processing. IEEE, 2017:379-382.
- [3] HU Y, SUN X, NIE X, et al. An Enhanced LSTM for Trend Following of Time Series [J]. IEEE Access, 2019, 7: 34020-34030.
- [4] ZHANG Q, LI F, LONG F, et al. Vehicle Emission Forecasting Based on Wavelet Transform and Long Short-Term Memory Network[J]. IEEE Access, 2018, 6:56984-56994.
- [5] WANG R, LIANG X, ZHU X, et al. A Feasibility of Respiration Prediction Based on Deep Bi-LSTM for Real-Time Tumor Tracking[J]. IEEE Access, 2018, 6:51262-51268.
- [6] SAAD W. Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural network[J]. IEEE Transactions on Neural Networks, 1998, 9(6):1456-70.
- [7] LIN S, RUNGER G C. GCRNN: Group-Constrained Convolutional Recurrent Neural Network [J]. IEEE Transactions on Neural Networks and Learning Systems, 2018, 29(10):4709-4718.
- [8] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8):1735-1780.
- [9] ZHANG Q, WANG H, DONG J, et al. Prediction of Sea Surface Temperature Using Long Short-Term Memory[J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(10):1745-1749.
- [10] YANG Y, DONG J, SUN X, et al. A CFCC-LSTM Model for Sea Surface Temperature Prediction[J]. IEEE Geoscience and Remote Sensing Letters, 2018, 15(2):207-211.
- [11] KARIM F, MAJUMDAR S, DARABI H, et al. LSTM Fully Convolutional Networks for Time Series Classification[J]. IEEE Access, 2017, 6(99):1662-1669.
- [12] LIU F, CAI M, WANG L, et al. An Ensemble Model Based on Adaptive Noise Reducer and Over-Fitting Prevention LSTM for Multivariate Time Series Forecasting[J]. IEEE Access, 2019, 6:26102-26115.
- [13] ZHANG S, WANG Y, LIU M, et al. Data-based Line Trip Fault Prediction in Power Systems Using LSTM Networks and SVM [J]. IEEE Access, 2017, 6:1-1.
- [14] CHO K, VAN MERRIËNBOER B, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[J]. arXiv:1406.1078, 2014.
- [15] CHEN W, ZHANG Y, YEO C K, et al. Stock market prediction using neural network through news on online social networks [C]//2017 International Smart Cities Conference (ISC2). IEEE, 2017:1-6.
- [16] SUN L, SU T, ZHOU S, et al. GMU: A Novel RNN Neuron and Its Application to Handwriting Recognition[C]//2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR). IEEE Computer Society, 2017.
- [17] HASIB M M, NAYAK T, HUANG Y. A hierarchical LSTM model with attention for modeling EEG non-stationarity for human decision prediction[C]//IEEE EMBS International Conference on Biomedical & Health Informatics (BHI). IEEE, 2018:104-107.
- [18] TANG X. Large-Scale Computing Systems Workload Prediction Using Parallel Improved LSTM Neural Network[J]. IEEE Access, 2019, 7:40525-40533.
- [19] LIU Y, DONG S, LU M, et al. LSTM based reserve prediction for bank outlets[J]. Tsinghua Science and Technology, 2018, 24(1):77-85.
- [20] SU P, DING X R, ZHANG Y T, et al. Long-term blood pressure prediction with deep recurrent neural networks[C]//2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI). IEEE, 2018:323-328.